

2012 年度 修士論文要旨

# ELIZA 的アプローチによって未知語を疑似獲得する音声対話システム

関西学院大学大学院理工学研究科  
情報科学専攻 川端研究室 高橋 勉

ユーザの発話を聞き、理解し、話し返すことができる人工知能的な機能を有したインターフェースである音声対話システムは音声認識技術の進歩に伴い、介護における話し相手や、車載情報サービス、観光情報案内システムなど多岐にわたり人間はコンピュータと対話するようになり、今後さまざまなアプリケーションなどへの展開が期待されている。しかし、コンピュータとの対話と人間同士の対話と比較すると、コンピュータの応答や対話の進行が不自然になることも多い。

通常音声認識システムでは語彙を限定して、その語彙内単語に強制的にマッチングを行うので、未登録語を検出することができない。これは、音声認識の性能を確保するうえでやむを得ない。その原因の一つとして円滑性（発話者の意図に沿い、対話が速やかに行われること）の欠如が報告されている。円滑性の欠如はこういったコンピュータの単語辞書内に未登録な単語、即ち未知語に遭遇した場合にその現象は著しい。

本研究では未知語を処理する音声対話システムの検討を行った。テキストベースの対話において、未知語の単語を抽出すると、それを記憶しておいて、何らかのタイミングでシステムの応答に組み込む ELIZA のアプローチを用いてテキストベースではなく音声対話システムでの未知語を疑似獲得する手法を提案した。未知語の疑似獲得については音声認識でユーザ発話中から未知語を判定するために本研究では CFG に基づく認識を行い、未知語と判定された音声区間を切り出し、記憶するためにビタビアライメントによる未知語区間の決定を行った。

本研究で提案した音声対話システムを評価するために成人男性 9 名に被験者実験を行った。アンケート項目から「楽しい、愉快な」や「リラックスできる」、「興味深い」などの項目に影響を与えており、面白さや興味を記していることがわかる。しかし、「発声が不自然」や「音声の品質が悪い」などの項目にも特徴が見られた。この原因の一つは獲得した未知語彙をシステムが応答音声として用いたとき、あらかじめ収録していた音声との違いを感じたためである。このことから未知語を疑似獲得する対話システムは音質などを考慮した補正法が必要であると考えられる。

未知語区間の補正の手法として無声音で終わる場合にビタビアライメントで判定された未知語区間に 5 フレームを加え未知語の切り出しを行った。この実験から未知語区間の補正を行っていないシステムとビタビアライメントで判定された未知語区間に 5 フレームを加え未知語の切り出しを行ったシステムとの比較を行ったとき未知語区間の補正を行った方が未知語の獲得に対しよい印象を与えていることが分かった。